10/14/99

Keratinocyte Derived Interferon

Field of the Invention

This application claims benefit of 35 U.S.C. section 119(e) based on copending U.S. Provisional Application Serial No: 60/093,643, filed July 21, 1998, incorporated herein by reference in its entirety.

The present invention relates to a novel human gene encoding a polypeptide which is a member of the interferon family. More specifically, isolated nucleic acid molecules are provided encoding a human polypeptide named "Keratinocyte Derived Interferon" or "KDI". KDI polypeptides are also provided, as are vectors, host cells and recombinant methods for producing the same. Also provided are diagnostic methods for detecting disorders related to the immune system, and therapeutic methods for treating disorders of the immune system. The invention further relates to screening methods for identifying agonists and antagonists of KDI.

Background of the Invention

Interferons (IFNs) are a well known family of cytokines secreted by a large variety of eukaryotic cells upon exposure to various stimuli. The interferons have been classified by their chemical and biological characteristics into four groups: IFN-alpha (leukocytes), IFN-beta (fibroblasts), IFN-gamma (lymphocytes), and IFN-omega (leukocytes). IFN-alpha and beta are known as Type I interferons; IFN-gamma is known as a Type-II or immune interferon. A single functional gene in the human genome codes for interferon omega (IFN-omega), a monomeric glycoprotein distantly related in structure to IFN-alpha and IFN-beta, but unrelated to IFN-gamma. IFN-omega is secreted by virus-infected leukocytes as a major component of human leukocyte interferon. The IFNs exhibit anti-viral, immunoregulatory, and antiproliferative activity. The clinical potential of interferons has been recognized, and will be summarized below.

Anti-viral: IFNs have been used clinically for anti-viral therapy, for example, in the treatment of AIDS (Lane, *Semin. Oncol.* 18:46-52 (Oct. 1991)), viral hepatitis including chronic hepatitis B, hepatitis C (Woo, M.H. and Brunakis, T.G., *Ann. Parmacother*, 31:330-337 (March 1997); Gibas, A.L., *Gastroenterologist*, 1:129-142 (June 1993)), hepatitis D, papilloma viruses (Levine, L.A. et al., *Urology* 47:553-557 (April 1996)), herpes (Ho, M., *Annu. Rev. Med.* 38:51-59 (1987)), viral encephalitis (Wintergerst et al., *Infection*, 20:207-212 (July 1992)), and in the prophylaxis of rhinitis and respiratory infections (Ho, M., *Annu. Rev. Med.* 38:51-59 (1987)).

Anti-parasitic: IFNs have been suggested for anti-parasite therapy, for example, IFN-gamma for treating *Cryptosporidium parvum* infection (Rehg, J.E., *J. Infect. Des.* 174:229-232 (July 1996)).

Anti-bacterial: IFNs have been used clinically for anti-bacterial therapy. For example, IFN-gamma has been used in the treatment of multidrug-resistant pulmonary tuberculosis (Condos, R. *et al., Lancet 349*:1513-1515 (1997)).

Anti-cancer: Interferon therapy has been used in the treatment of numerous cancers (*e.g.*, hairy cell leukemia (Hofmann *et al., Cancer Treat. Rev. 12 (Suppl. B)*:33-37 (Dec. 1985)), acute myeloid leukemia (Stone, R.M. *et al. Am. J. Clin. Oncol. 16:*159-163 (April 1993)), osteosarcoma (Strander, H. *et al.*, Acta Oncol. 34:877-880 (1995)), basal cell carcinoma (Dogan, B. *et al., Cancer Lett. 91*:215-219 (May 1995)), glioma (Fetell, M.R. *et al., Cancer 65*: 78-83 (Jan. 1990)), renal cell carcinoma (Aso, Y. *et al. Prog. Clin. Biol. Res. 303*:653-659 (1989)), multiple myeloma (Peest, D. *et al., Br. J. Haematol. 94*:425-432 (Sept. 1996)), melanoma (Ikic, D. *et al., Int. J. Dermatol. 34*:872-874 (Dec. 1995)), and Hodgkin's disease (Rybak, M.E. *et al., J. Biol. Response Mod. 9*:1-4 (Feb. 1990)). Synergistic treatment of advanced cancer with a combination of alpha interferon and temozolomide has also been reported (Patent publication WO 9712630 to Dugan, M.H.).

Immunotherapy: IFNs have been used clinically for immunotherapy or more particularly, (1) for example, to prevent graft vs. host rejection, or to curtail the progresion of autoimmune diseases, such as arthritis, multiple sclerosis, (2) or diabetes (3). IFN-beta is approved of sale in the United States for the treatment (i.e., as an immunosuppressant) of multiple sclerosis. Recently it has been reported that patients with multiple sclerosis have diminished production of type I interferons and interleukin-2 (Wandinger, K.P. *et al., J. Neurol. Sci. 149*: 87-93 (1997)). In addition, immunotherapy with recombinant IFN-alpha (in combination with recombinant human IL-2) has been used successfully in lymphoma patients following autologous bone marrow or blood stem cell transplantation, that may intensify remission following translation (Nagler, A. *et al., Blood 89*: 3951-3959 (June 1997)).

Anti-allergy: The administration of IFN-gamma has been used in the treatment of allergies in mammals (See, Patent Publication WO 8701288 to Parkin, J.M. and Pinching, A.J.). It has also recently been demonstrated that there is a reduced production of IL-12 and IL-12-dependent IFN-gamma release in patients with allergic asthma (van der Pouw Kraan, T.C. *et al., J. Immunol. 158:*5560-5565 (1997)). Thus, IFN may be useful in the treatment of allergy by inhibiting the humoral response.

Vaccine adjuvantation: Interferons may be used as an adjuvant or coadjuvant to enhance or simulate the immune response in cases of prophylactic or therapeutic vaccination (Heath, A.W. and Playfair, J.H.L., *Vaccine 10*:427-434 (1992)).

Clearly, there exists a need in the art for the discovery of novel interferon proteins for numerous applications, in *e.g.*, immunotherapy, as well as anti-viral, anti-parasitic, anti-bacterial, or anti-cancer therapies, or any medical condition or situation where increased interferon activity is desired.

Summary of the Invention

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding at least a portion of the KDI polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 or the complete amino acid sequence encoded by the cDNA clone deposited as plasmid DNA as ATCC Deposit Number 203500 on December 1, 1998. The nucleotide sequence determined by sequencing the deposited KDI clone (HKAPI15) which is shown in Figure 1 (SEQ ID NO:1), contains an open reading frame encoding a complete polypeptide of 207 amino acid residues, including an initiation codon encoding an N-terminal methionine at nucleotide positions 35-37. Nucleic acid molecules of the invention include those encoding the complete amino acid sequence excepting the N-terminal methionine shown in SEQ ID NO:2, which molecules also can encode additional amino acids fused to the N-terminus of the KDI amino acid sequence.

The encoded polypeptide has a predicted leader sequence of 27 amino acids underlined in Figure 1; and the amino acid sequence of the predicted mature KDI protein is also shown in Figure 1 as amino acid residues 28-207 and as residues 1-207 in SEQ ID NO:2.

Residues 165 to 183 of Figure 1 (SEQ ID NO:2) comprise the signature sequence for interferon polypeptides, and the KDI polypeptide in particular. Thus, preferred are polypeptides comprising residues 165 to 183 of Figure 1 as are polynucleotides encoding such polypeptides.

Thus, one aspect of the invention provides an isolated nucleic acid molecule comprising a polynucleotide comprising a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding the KDI polypeptide having the complete amino acid sequence in SEQ ID NO:2; (b) a nucleotide sequence encoding the KDI polypeptide having the complete amino acid sequence in SEQ ID NO:2 excepting the N-terminal methionine (i.e., residues 2-207 of SEQ ID NO:2); (c) a nucleotide sequence encoding the mature KDI polypeptide shown as residues 28-207 in SEQ ID NO:2; (d) a nucleotide sequence encoding residues 165-183 of SEQ ID NO:2; (e) a nucleotide sequence encoding the complete polypeptide encoded by the human cDNA contained in clone HKAPI15; (f) a nucleotide sequence encoding the complete

polypeptide encoded by the human cDNA contained in clone HKAPI15 excepting the N-terminal methionine; (g) a nucleotide sequence encoding the mature polypeptide encoded by the human cDNA contained in clone HKAPI15; and (h) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f) or (g) above.

Further embodiments of the invention include isolated nucleic acid molecules that comprise a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical, to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g) or (h), above, or a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide in (a), (b), (c), (d), (e), (f), (g) or (h), above. This polynucleotide which hybridizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues. An additional nucleic acid embodiment of the invention relates to an isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of a KDI polypeptide having an amino acid sequence in (a), (b), (c), (d), (e), (f) or (g), above.

The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of KDI polypeptides or peptides by recombinant techniques.

The invention further provides an isolated KDI polypeptide comprising an amino acid sequence selected from the group consisting of: (a) the amino acid sequence of the full-length KDI polypeptide having the complete amino acid sequence shown in SEQ ID NO:2; (b) the amino acid sequence of the full-length KDI polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 excepting the N-terminal methionine (i.e., residues 2 to 207 of SEQ ID NO:2); the amino acid sequence of the mature KDI polypeptide shown as residues 28-207 in SEQ ID NO:2; (d) the amino acid sequence shown as residues 165 to 183 of SEQ ID NO:2; (e) the full length KDI polypeptide encoded by the human cDNA contained in clone HKAPI15; (f) the full-length KDI polypeptide encoded by the human cDNA contained in clone HKAPI15 excepting the N-terminal methionine; and (g) the mature KDI polypeptide encoded by the human cDNA contained in clone HKAPI15. The polypeptides of the present invention also include polypeptides having an amino acid sequence at least 80% identical, more preferably at least 90% identical, and still more preferably 95%, 96%, 97%, 98% or 99% identical to those described in (a), (b), (c), (d), (e), (f) or (g) above, as well as polypeptides having an amino acid sequence with at least 90% similarity, and more preferably at least 95% similarity, to those above.

An additional embodiment of this aspect of the invention relates to a peptide or polypeptide which comprises the amino acid sequence of an epitope-bearing portion of a KDI polypeptide having an amino acid sequence described in (a), (b), (c), (d), (e), (f), or (g),

above. Peptides or polypeptides having the amino acid sequence of an epitope-bearing portion of a KDI polypeptide of the invention include portions of such polypeptides with at least six or seven, preferably at least nine, and more preferably at least about 30 amino acids to about 50 amino acids, although epitope-bearing polypeptides of any length up to and including the entire amino acid sequence of a polypeptide of the invention described above also are included in the invention.

In another embodiment, the invention provides an isolated antibody that binds specifically to a KDI polypeptide having an amino acid sequence described in (a), (b), (c), (d), (e), (f) or (g) above. The invention further provides methods for isolating antibodies that bind specifically to a KDI polypeptide having an amino acid sequence as described herein. Such antibodies are useful therapeutically as described below.

The invention also provides for pharmaceutical compositions comprising KDI polypeptides which may be employed, for instance, to treat immune system-related disorders such as viral infection, parasitic infection, bacterial infection, cancer, autoimmune disease, multiple sclerosis, lymphoma and allergy. Methods of treating individuals in need of interferon polypeptides are also provided.

The invention further provides compositions comprising a KDI polynucleotide or a KDI polypeptide for administration to cells *in vitro*, to cells *ex vivo* and to cells *in vivo*, or to a multicellular organism. In certain particularly preferred embodiments of this aspect of the invention, the compositions comprise a KDI polynucleotide for expression of a KDI polypeptide in a host organism for treatment of disease. Particularly preferred in this regard is expression in a human patient for treatment of a dysfunction associated with aberrant endogenous activity of an interferon.

The present invention also provides a screening method for identifying compounds capable of enhancing or inhibiting a biological activity of the KDI polypeptide, which involves contacting a receptor which is enhanced by the KDI polypeptide with the candidate compound in the presence of a KDI polypeptide, assaying, for example, anti-viral activity in the presence of the candidate compound and the KDI polypeptide, and comparing the activity to a standard level of activity, the standard being assayed when contact is made between the receptor and KDI in the absence of the candidate compound. In this assay, an increase in activity over the standard indicates that the candidate compound is an agonist of KDI activity and a decrease in activity compared to the standard indicates that the compound is an antagonist of KDI activity.

It has been discovered that KDI is expressed in keratinocytes. Therefore, nucleic acids of the invention are useful as hybridization probes for differential identification of the tissue(s) or cell type(s) present in a biological sample. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). In addition, for a number of disorders of the above tissues or cells, particularly of the immnue system, significantly higher or lower levels of KDI gene expression

may be detected in certain tissues (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to a "standard" KDI gene expression level, i.e., the KDI expression level in healthy tissue from an individual not having the immune system disorder. Thus, the invention provides a diagnostic method useful during diagnosis of such a disorder, which involves: (a) assaying KDI gene expression level in cells or body fluid of an individual; (b) comparing the KDI gene expression level with a standard KDI gene expression level, whereby an increase or decrease in the assayed KDI gene expression level compared to the standard expression level is indicative of disorder in the immune system.

An additional aspect of the invention is related to a method for treating an individual in need of an increased level of interferon activity in the body comprising administering to such an individual a composition comprising a therapeutically effective amount of an isolated KDI polypeptide of the invention or an agonist thereof.

A still further aspect of the invention is related to a method for treating an individual in need of a decreased level of interferon activity in the body comprising, administering to such an individual a composition comprising a therapeutically effective amount of a KDI antagonist. Preferred antagonists for use in the present invention are KDI-specific antibodies.

Brief Description of the Figures

Figure 1 shows the nucleotide sequence (SEQ ID NO:1) and amino acid sequence (SEQ ID NO:2) of KDI.

Figure 2 shows the regions of identity between the amino acid sequences of the KDI protein and translation product of the human mRNA for Interferon Omega (SEQ ID NO:3), determined by the computer program Bestfit (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711).

Figure 3 shows an analysis of the KDI amino acid sequence. Alpha, beta, turn and coil regions; hydrophilicity and hydrophobicity; amphipathic regions; flexible regions; antigenic index and surface probability are shown. In the "Antigenic Index - Jameson-Wolf" graph, the positive peaks indicate locations of the highly antigenic regions of the KDI protein, i.e., regions from which epitope-bearing peptides of the invention can be obtained.

Figure 4 shows an alignment of the KDI polypeptide of the present invention with several other members of the interferon polypeptide family. Shown is human interferon beta-1 (SEQ ID NO:4), human placental interferon (SEQ ID NO:5), human interferon omega (SEQ ID NO:6), human interferon alpha-c (SEQ ID NO:7), human interferon alpha-F (SEQ ID NO:8), human interferon II-1 (SEQ ID NO:9), human alpha interferon-N (SEQ ID NO:10), bovine TP-1 (SEQ ID NO:11), ovi TP-1 (SEQ ID NO:12), pig TP (SEQ ID NO:13), human interferon beta 2a (IL-6) (SEQ ID NO:14), bovine interferon beta-2 (SEQ ID NO:15), bovine interferon beta-1

(SEQ ID NO:20), synthetic interferon beta-1 (SEQ ID NO:21). The alignment was produced by the Megalign routine using the Clustal method with PAM250 residue weight table. Megalign is contained within the DNAstar suite of programs.

Detailed Description

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding a Keratinocyte-Derived Interferon polypeptide (hereinafter "KDI") having the amino acid sequence shown in SEQ ID NO:2. The nucleotide sequence shown in Figure 1 (SEQ ID NO:1) was obtained by sequencing the human HKAPI15 cDNA clone which was deposited on December 1, 1998 at the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209, USA, and given accession number ATCC 203500. The deposited cDNA is contained in the plasmid pCMVSport 2.0 (Life Technologies, Gaithersburg MD) and can be excised by the SalI/NotI restriction enzyme sites flanking the human cDNA.

The KDI protein of the present invention shares sequence homology with many members of the interferon family, noteably the translation product of the human mRNA for IFN-omega (Figure 2) (SEQ ID NO:3). IFN-omega has been shown to inhibit the proliferation of a variety of tumor cell lines in vitro, stimulate natural killer cell activity, enhance expression of major histocompatibility complex class I (but not class II) antigens and inhibit proliferation of lymphocytes stimulated with mitogens or allogeneic cells. Adolf, G.R., Human interferon omega--a review, *Mult Scler* 1995;1 Suppl 1:S44-S47. The expression of KDI has been determined to be primarily found in keratinocytes, dendritic cells and monocytes, but is particularly stong in keratinocytes. Stimulation of keratinocytes with TNF-α or PolyIC (simulating viral infection) specifically and rapidly stimulates overexpression of the KDI transcript. Based on its structural similarity to IFN-omega and its increased expression in response to simulated viral infection, KDI is believed to share many of its biological activities of INF-Omega and other interferon proteins, including, inhibition of tumor proliferation, antiviral activities, NK cell activiation, and immune system enhancement.

Nucleic Acid Molecules

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer (such as the Model 373 from Applied Biosystems, Inc., Foster City, CA), and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more

typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

By "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U), where each thymidine deoxyribonucleotide (T) in the specified deoxyribonucleotide sequence is replaced by the ribonucleotide uridine (U).

Using the information provided herein, such as the nucleotide sequence in Figure 1 (SEQ ID NO:1), a nucleic acid molecule of the present invention encoding a KDI polypeptide may be obtained using standard molecular biology procedures, such as those for cloning cDNAs using mRNA as starting material. Illustrative of the invention, the nucleic acid molecule described in Figure 1 (SEQ ID NO:1) was discovered in a cDNA library derived from isolated keratinocytes.

The nucleotide sequence of the KDI DNA of Figure 1 (SEQ ID NO:1) contains an open reading frame encoding a protein of 207 amino acid residues, with an initiation codon at nucleotide positions 35-37 of the nucleotide sequence in Figure 1 (SEQ ID NO:1). The amino acid sequence of the KDI protein shown in SEQ ID NO:2 is about 35% identical to IFN-omega, (Figure 2). The sequences of INF-Omega can be accessed through GenBank with Accession No. gb|A12140.

As one of ordinary skill would appreciate, due to the possibilities of sequencing errors discussed above, the actual complete KDI polypeptide encoded by the deposited cDNA, which comprises about 207 amino acids, may be somewhat longer or shorter. More generally, the actual open reading frame may be anywhere in the range of ±20 amino acids, more likely in the range of ±10 amino acids, of that predicted from the methionine codon at the N-terminus shown in Figure 1 (SEQ ID NO:1).

Leader and Mature Sequences

The amino acid sequence of the complete KDI protein includes a leader sequence and a mature protein, as shown in SEQ ID NO:2. More in particular, the present invention provides nucleic acid molecules encoding a mature form of the KDI protein. Thus, according to the signal hypothesis, once export of the growing protein chain across the rough endoplasmic

reticulum has been initiated, proteins secreted by mammalian cells have a signal or secretory leader sequence which is cleaved from the complete polypeptide to produce a secreted "mature" form of the protein. Most mammalian cells and even insect cells cleave secreted proteins with the same specificity. However, in some cases, cleavage of a secreted protein is not entirely uniform, which results in two or more mature species of the protein. Further, it has long been known that the cleavage specificity of a secreted protein is ultimately determined by the primary structure of the complete protein, that is, it is inherent in the amino acid sequence of the polypeptide. Therefore, the present invention provides a nucleotide sequence encoding the mature KDI polypeptide having the amino acid sequence encoded by the human cDNA in clone HKAPI15 (ATCC Deposit No. 203500). By the "mature KDI polypeptide having the amino acid sequence encoded by the human cDNA in clone HKAPI15" is meant the mature form(s) of the KDI protein produced by expression in a mammalian cell (e.g., COS cells, as described below) of the complete open reading frame encoded by the human DNA sequence of the clone contained in the deposited vector.

In addition, methods for predicting whether a protein has a secretory leader as well as the cleavage point for that leader sequence are available. For instance, the method of McGeoch (*Virus Res. 3*:271-286 (1985)) uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje (*Nucleic Acids Res. 14*:4683-4690 (1986)) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2 where +1 indicates the amino terminus of the mature protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80% (von Heinje, *supra*). However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the complete KDI polypeptide was analyzed by a computer program "PSORT", available from Dr. Kenta Nakai of the Institute for Chemical Research, Kyoto University (see K. Nakai and M. Kanehisa, *Genomics 14:*897-911 (1992)), which is an expert system for predicting the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated." The computation analysis above predicted one potential cleavage site within the complete amino acid sequence shown in SEQ ID NO:2; that is, between residues 27 and 28 in Figure 1 (SEQ ID NO:2). Of course the exact address of the cleavage site used by naturally occurring enzymes may vary slightly from the predicted clevage site and may vary between species. Accordingly, mature polypeptides beginning from about residue 20 to about residue 34 are provided. More in particular, the invention provides a polypeptide having a portion of SEQ ID NO:2 as follows: residues 20-207 in SEQ ID NO:2, residues 21-207 in SEQ ID NO:2, residues 22-207 in SEQ ID NO:2, residues 23-207 in SEQ ID NO:2, residues 24-207 in SEQ ID NO:2, residues 25-207 in SEQ ID NO:2, residues 26-207

in SEQ ID NO:2, residues 27-207 in SEQ ID NO:2, residues 28-207 in SEQ ID NO:2, residues 29-207 in SEQ ID NO:2, residues 30-207 in SEQ ID NO:2, residues 31-207 in SEQ ID NO:2, residues 32-207 in SEQ ID NO:2, residues 33-207 in SEQ ID NO:2, and residues 34-127 in SEQ ID NO:2. The invention also provides polynucleotides encoding such polypeptides.

As indicated, nucleic acid molecules of the present invention may be in the form of RNA, or in the form of DNA. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

In specific embodiments, the polynucleotides of this invention are less than 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb or 7.5 kb in length. In a further embodiment, polynucleotides of the invention comprise at least 15 contiguous nucleotides of KDI coding sequence, but do not comprise all or a portion of any KDI intron. In another embodiment, the nucleic acid comprising KDI coding sequence does not contain coding sequences of a genomic flanking gene (i.e. 5' or 3' to the KDI coding sequence in the genome).

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. However, a nucleic acid contained in a clone that is a member of a library (e.g., a genomic or cDNA library) that has not been isolated from other members of the library (e.g., in the form of a homogenous solution containing the clone and other members of the library) or a chromosome isolated or removed from a cell or cell lysate (e.g., a "chromosome spread", as in a karyotype), is not "isolated" for the purposes of this invention. As discussed further herein, isolated nucleic acid molecules according to the present invention may be produced naturally, recombinantly or synthetically.

Isolated nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) with an initiation codon at positions 35-37 of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1).

Also included are DNA molecules comprising the coding sequence for the KDI protein lacking an N-terminal methionine shown at positions 2-207 of SEQ ID NO:2.

In addition, isolated nucleic acid molecules of the invention include DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode the KDI protein. Of course, the genetic code and species-specific codon preferences are well known in the art. Thus, it would be routine for one skilled in the art to generate the degenerate variants described above, for instance, to

optimize codon expression for a particular host (e.g., change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

In another aspect, the invention provides isolated nucleic acid molecules encoding the KDI polypeptide having an amino acid sequence encoded by the human cDNA in clone HKAPI15 (ATCC Deposit No. 203500). Preferably, this nucleic acid molecule will encode the mature polypeptide encoded by the above-described deposited human cDNA.

The invention further provides an isolated nucleic acid molecule having the nucleotide sequence shown in Figure 1 (SEQ ID NO:1) or the nucleotide sequence of the KDI cDNA contained in the above-described deposited clone, or a nucleic acid molecule having a sequence complementary to one of the above sequences. Such isolated molecules, particularly DNA molecules, are useful for production of the KDI polypeptide of the invention and as a probe for detection of mRNA in cells transfected with a vector for the purpose of producing KDI; i.e., as a marker for determining expression of the heterologous gene in a host cell.

The present invention is further directed to nucleic acid molecules encoding portions of the nucleotide sequences described herein as well as to fragments of the isolated nucleic acid molecules described herein. In particular, the invention provides a polynucleotide having a nucleotide sequence representing the portion of SEQ ID NO:1 which consists of positions 35-655 of SEQ ID NO:1. Other particularly preferred polynucleotide fragments of the invention comprise, or alternatively, consist of nucleotide residues 38-655, 41-655, 44-655, 47-655, 50-655, 53-655, 56-655, 59-655, 62-655, 65-655, 68-655, 71-655, 74-655, 77-655, 80-655, 83-655, 86-655, 89-655, 92-655, 95-655, 98-655, 101-655, 104-655, 107-655, 110-655, 113-655, 116-655, 119-655, 122-655, 125-655, 128-655, 131-655, 134-655, 137-655, 140-655, 143-655, 146-655, 149-655, 152-655, 155-655, 158-655, 161-655, 164-655, 167-655, 170-655, 173-655, 176-655, 179-655, 182-655, 185-655, 188-655, 191-655, 194-655, 197-655, 200-655, 203-655, 206-655, 209-655, 212-655, 215-655, 218-655, 221-655, 224-655, 227-655, 230-655, 233-655, 236-655, 239-655, 242-655, 245-655, 248-655, 251-655, 254-655, 257-655, 260-655, 263-655, 266-655, 269-655, 272-655, 275-655, 278-655, 281-655, 284-655, 287-655, 290-655, 293-655, 296-655, 299-655, 302-655, 305-655, 308-655, 311-655, 314-655, 317-655, 320-655, 323-655, 326-655, 329-655, 332-655, 335-655, 338-655, 341-655, 344-655, 347-655, 350-655, 353-655, 356-655, 359-655, 362-655, 365-655, 368-655, 371-655, 374-655, 377-655, 380-655, 383-655, 386-655, 389-655, 392-655, 395-655, 398-655, 401-655, 404-655, 407-655, 410-655, 413-655, 416-655, 419-655, 422-655, 425-655, 428-655, 431-655, 434-655, 437-655, 440-655, 443-655, 446-655, 449-655, 452-655, 455-655, 458-655, 461-655, 464-655, 467-655, 470-655, 473-655, 476-655, 479-655, 482-655, 485-655, 488-655, 491-655, 494-655, 497-655, 500-655, 503-655, 506-655, 509-655, 512-655, 515-655, 518-655, 521-655, 524-655, 527-655, 530-655, 533-655, 536-655, 539-655, 542-655, 545-655, 548-655, 551-655, 554-655, 557-655, 560-655, 563-655, 566-655, 569-655, 572-655, 575-655, 578-655, 581-655, 584-655, 587-655, 590-655, 593-655, 596-655,

599-655, 602-655, 605-655, 608-655, 611-655, 614-655, 617-655, 620-655, 623-655, 626-655, 629-655, 632-655, and 635-655 of SEQ ID NO:1. Still other particularly preferred polynucleotide fragments of the invention comprise, or alternatively, consist of nucleotide residues 38-68, 38-71, 38-74, 38-77, 38-80, 38-83, 38-86, 38-89, 38-92, 38-95, 38-98, 38-101, 38-104, 38-107, 38-110, 38-113, 38-116, 38-119, 38-122, 38-125, 38-128, 38-131, 38-134, 38-137, 38-140, 38-143, 38-146, 38-149, 38-152, 38-155, 38-158, 38-161, 38-164, 38-167, 38-170, 38-173, 38-176, 38-179, 38-182, 38-185, 38-188, 38-191, 38-194, 38-197, 38-200, 38-203, 38-206, 38-209, 38-212, 38-215, 38-218, 38-221, 38-224, 38-227, 38-230, 38-233, 38-236, 38-239, 38-242, 38-245, 38-248, 38-251, 38-254, 38-257, 38-260, 38-263, 38-266, 38-269, 38-272, 38-275, 38-278, 38-281, 38-284, 38-287, 38-290, 38-293, 38-296, 38-299, 38-302, 38-305, 38-308, 38-311, 38-314, 38-317, 38-320, 38-323, 38-326, 38-329, 38-335, 38-338, 38-341, 38-344, 38-347, 38-350, 38-353, 38-356, 38-359, 38-362, 38-365, 38-368, 38-371, 38-374, 38-377, 38-380, 38-383, 38-386, 38-389, 38-392, 38-395, 38-398, 38-401, 38-404, 38-407, 38-410, 38-413, 38-416, 38-419, 38-422, 38-425, 38-428, 38-431, 38-434, 38-437, 38-440, 38-443, 38-446, 38-449, 38-452, 38-455, 38-458, 38-461, 38-464, 38-467, 38-470, 38-473, 38-476, 38-479, 38-482, 38-485, 38-488, 38-491, 38-494, 38-497, 38-500, 38-503, 38-506, 38-509, 38-512, 38-515, 38-518, 38-521, 38-524, 38-527, 38-530, 38-533, 38-536, 38-539, 38-542, 38-545, 38-548, 38-551, 38-554, 38-557, 38-560, 38-563, 38-566, 38-569, 38-572, 38-575, 38-578, 38-581, 38-584, 38-587, 38-590, 38-593, 38-596, 38-599, 38-602, 38-605, 38-608, 38-611, 38-614, 38-617, 38-620, 38-623, 38-626, 38-629, 38-632, and 38-635 of SEQ ID NO:1.

Further, the invention includes a polynucleotide comprising any portion of at least about 30 contiguous nucleotides, preferably at least about 50 contiguous nucleotides, of SEQ ID NO:1.

More generally, by a fragment of an isolated nucleic acid molecule having the nucleotide sequence of the deposited cDNA or the nucleotide sequence shown in Figure 1 (SEQ ID NO:1) is intended fragments at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length which are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments 50-600 nt in length (fragments of 400 nt, 450 nt, 500 nt, 550 nt and 600 nt in length are specifically contempleted as are fragments of all lengths between 15 and 600 but will not be specifically recited for space considerations) are also useful according to the present invention as are fragments corresponding to most, if not all, of the nucleotide sequence of the deposited cDNA or as shown in Figure 1 (SEQ ID NO:1). By a fragment at least 20 nt in length, for example, is intended fragments which include 20 or more contiguous bases from the nucleotide sequence of the deposited cDNA or the nucleotide sequence as shown in Figure 1 (SEQ ID NO:1) and may, of course, comprise additional nucleic acid sequences not derived from SEQ ID NO:1 (or the deposited cDNA) fused to either end of the 20+ contiguous bases from SEQ

ID NO:1 or the deposited cDNA. Preferred nucleic acid fragments of the present invention include nucleic acid molecules encoding epitope-bearing portions of the KDI polypeptide as identified in Figure 3 and described in more detail below.

In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention described above, for instance, the human cDNA in clone HKAPI15 (ATCC Deposit No. 203500). By "stringent hybridization conditions" is intended overnight incubation at 42° C in a solution comprising: 50% formamide, 5x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters twice in 0.1x SSC at about 65° C.

By a polynucleotide which hybridizes to a "portion" of a polynucleotide is intended a polynucleotide (either DNA or RNA) hybridizing to at least about 15 nucleotides (nt), and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably about 30-70 (e.g., 50) nt of the reference polynucleotide. These are useful as diagnostic probes and primers as discussed above and in more detail below.

By a portion of a polynucleotide of "at least 20 nt in length," for example, is intended 20 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide (e.g., the deposited cDNA or the nucleotide sequence as shown in Figure 1 (SEQ ID NO:1)). Of course, a polynucleotide which hybridizes only to a poly A sequence (such as the 3' terminal poly(A) tract of the KDI cDNA shown in Figure 1 (SEQ ID NO:1)), or to a complementary stretch of T (or U) residues, would not be included in a polynucleotide of the invention used to hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

In specific embodiments, the polynucleotides of the invention are less than 100000 kb, 50000 kb, 10000 kb, 1000 kb, 500, kb, 400 kb, 350 kb, 300 kb, 250 kb, 200 kb, 175 kb, 150 kb, 125 kb, 100 kb, 75 kb, 50 kb, 40 kb, 30 kb, 25 kb, 20 kb, 15 kb, 7.5 kb, or 5 kb in length.

As indicated, nucleic acid molecules of the present invention which encode a KDI polypeptide may include, but are not limited to those encoding the amino acid sequence of the complete polypeptide, by itself; and the coding sequence for the complete polypeptide and additional sequences, such as those encoding an added secretory leader sequence, such as a pre-, or pro- or prepro- protein sequence.

Also encoded by nucleic acids of the invention are the above protein sequences together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing, including splicing and polyadenylation signals, for

example - ribosome binding and stability of mRNA; an additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities.

Thus, the sequence encoding the polypeptide may be fused to a marker sequence, such as a sequence encoding a peptide which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz *et al., Proc. Natl. Acad. Sci. USA 86*:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. The "HA" tag is another peptide useful for purification which corresponds to an epitope derived from the influenza hemagglutinin protein, which has been described by Wilson *et al., Cell 37*: 767 (1984). As discussed below, other such fusion proteins include the KDI fused to Fc at the N- or C-terminus.

Variant and Mutant Polynucleotides

The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the KDI protein. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. *Genes II*, Lewin, B., ed., John Wiley & Sons, New York (1985). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

Such variants include those produced by nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the KDI protein or portions thereof. Also especially preferred in this regard are conservative substitutions.

Further embodiments include an isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical to a polynucleotide selected from the group consisting of: (a) a nucleotide sequence encoding the KDI polypeptide having the complete amino acid sequence in SEQ ID NO:2; (b) a nucleotide sequence encoding the KDI polypeptide having the complete amino acid sequence in SEQ ID NO:2 excepting the N-terminal methionine (i.e., residues 2-161 of SEQ ID NO:2; (c) a nucleotide sequence encoding the mature KDI polypeptide having the sequence shown as residues 28-207 in SEQ ID NO:2; (d) a nucleotide sequence encoding residues 165 to 183 shown in SEQ ID NO:2; (e) a nucleotide sequence encoding the complete amino acid sequence encoded by the human cDNA contained in clone

HKAPI15; (f) a nucleotide sequence encoding the complete amino acid sequence encoded by the human cDNA contained in clone HKAPI15 excepting the N-terminal methionine; (g) a nucleotide sequence encoding the amino acid sequence of the mature polypeptide encoded by the human cDNA contained in clone HKAPI15; and (h) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g) and (h), above.

Further embodiments of the invention include isolated nucleic acid molecules that comprise a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical, to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g) or (h), above, or a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide in (a), (b), (c), (d), (e), (f), (g) or (h) above. This polynucleotide which hybridizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues. An additional nucleic acid embodiment of the invention relates to an isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of a KDI polypeptide having an amino acid sequence in (a), (b), (c), (d), (e), (f) or (g), above.

The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of KDI polypeptides or peptides by recombinant techniques.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a KDI polypeptide is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the KDI polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular nucleic acid molecule is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the nucleotide sequence shown in Figure 1 or to the nucleotides sequence of the deposited cDNA clone can be determined conventionally using known computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575

Science Drive, Madison, WI 53711). Bestfit uses the local homology algorithm of Smith and Waterman to find the best segment of homology between two sequences (*Advances in Applied Mathematics* **2**:482-489 (1981)). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identiy are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the lenght of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base

subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The present application is directed to nucleic acid molecules at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1) or to the nucleic acid sequence of the deposited DNA, irrespective of whether they encode a polypeptide having KDI activity. This is because even where a particular nucleic acid molecule does not encode a polypeptide having KDI activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having KDI activity include, inter alia, isolating allelic variants in a cDNA library.

Preferred, however, are nucleic acid molecules having sequences at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1) or to the nucleic acid sequence of the deposited DNA which do, in fact, encode a polypeptide having KDI protein activity. By "a polypeptide having KDI activity" is intended polypeptides exhibiting activity similar, but not necessarily identical, to an activity of the KDI protein of the invention, as measured in a particular biological assay. For example, the KDI protein of the present invention inhibits bone marrow colony formation in-vitro and can be assayed according to the method of Tiefenthaler M. *et al.* (*Interferon Cytokine Res,* 1997 Jun;17(6):327-329, incorporated herein by reference in its entirety). In addition, KDI can inhibit GM-CSF induced proliferation of the erythroleukaemic cell line TF-1 according to the assays reported by Mire-Sluis A.R. et al. (*J. Immunol. Methods* 1996 Sep 9;195(1-2):55-61, incorporated herein by reference). Alternatively, KDI can be assayed for classical anti-viral activity by any of several assays known to those of skill in the art, for example, in the assay reported by Sugiyama, K. et al. (*Yakugaku Zasshi* 1995 May;115(5):390-393).

The KDI protein of the present invention inhibits bone marrow proliferation and shows anti-viral activity in a dose-dependent manner in the above-described assays. Thus, “a polypeptide having KDI protein activity” includes polypeptides that also exhibit any of the same activities in the above-described assays in a dose-dependent manner. Although the degree of dose-dependent activity need not be identical to that of the KDI protein, preferably, “a polypeptide having KDI protein activity” will exhibit substantially similar dose-dependence in a given activity as compared to the KDI protein (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity relative to the reference KDI protein).

Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 90%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequence of the deposited cDNA or the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1) will encode a polypeptide "having KDI protein activity." In fact, since degenerate variants of these nucleotide sequences all encode the same polypeptide, this will be clear to the skilled artisan even without performing the above described comparison assay. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having KDI protein activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid), as further described below.

Vectors and Host Cells

The present invention also relates to vectors which include the isolated DNA molecules of the present invention, host cells which are genetically engineered with the recombinant vectors, and the production of KDI polypeptides or fragments thereof by recombinant techniques. The vector may be, for example, a phage, plasmid, viral or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac*, *trp*, *phoA* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, Streptomyces and *Salmonella typhimurium* cells;

fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293 and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc., *supra*; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis *et al., Basic Methods In Molecular Biology* (1986).

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art. A preferred fusion protein comprises a heterologous region from immunoglobulin that is useful to stabilize and purify proteins. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is thoroughly advantageous for use in therapy and diagnosis and thus results, for example, in improved pharmacokinetic properties (EP-A 0232 262). On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified in the advantageous manner described. This is the case when Fc portion proves to be a hindrance to use in therapy and diagnosis, for example when the fusion protein is to be used as antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. See, D. Bennett *et al., J. Molecular Recognition 8*:52-58 (1995) and K. Johanson *et al., J. Biol. Chem. 270*:9459-9471 (1995).

The KDI protein can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction,

anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Polypeptides of the present invention include: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Polypeptides and Fragments

The invention further provides an isolated KDI polypeptide having the amino acid sequence encoded by the deposited DNA, or the amino acid sequence in SEQ ID NO:2, or a peptide or polypeptide comprising a portion of the above polypeptides.

Variant and Mutant Polypeptides

To improve or alter the characteristics of KDI polypeptides, protein engineering may be employed. Recombinant DNA technology known to those skilled in the art can be used to create novel mutant proteins or "muteins including single or multiple amino acid substitutions, deletions, additions or fusion proteins. Such modified polypeptides can show, e.g., enhanced activity or increased stability. In addition, they may be purified in higher yields and show better solubility than the corresponding natural polypeptide, at least under certain purification and storage conditions.

N-Terminal and C-Terminal Deletion Mutants

For instance, for many proteins, including the extracellular domain of a membrane associated protein or the mature form(s) of a secreted protein, it is known in the art that one or more amino acids may be deleted from the N-terminus or C-terminus without substantial loss of biological function. For instance, Ron et al., *J. Biol. Chem., 268*:2984-2988 (1993) reported modified KGF proteins that had heparin binding activity even if 3, 8, or 27 amino-terminal amino acid residues were missing. In the present case, since the protein of the invention is a member of the interferon polypeptide family, deletions of N-terminal amino acids up to the

cysteine at position 59 as shown in SEQ ID NO:2 may retain some biological activity such as antiviral activity or inhibition of bone marrow proliferation. Polypeptides having further N-terminal deletions including the Cys-59 residue in SEQ ID NO:2 would not be expected to retain such biological activities because it is known that this residue in an interferon-related polypeptide is conserved among many, if not all, members of the family as is Leucine residue immediately adjacent to it (residue 60). The cysteine residue at position 59 is thought to be required for forming a disulfide bridge to provide structural stability which is needed for receptor binding and signal transduction.

However, even if deletion of one or more amino acids from the N-terminus of a protein results in modification of loss of one or more biological functions of the protein, other biological activities may still be retained. Thus, the ability of the shortened protein to induce and/or bind to antibodies which recognize the complete protein generally will be retained when less than the majority of the residues of the complete protein are removed from the N-terminus. Whether a particular polypeptide lacking N-terminal residues of a complete protein retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art.

Accordingly, the present invention further provides polypeptides having one or more residues deleted from the amino terminus of the amino acid sequence of the KDI shown in SEQ ID NO:2, up to the Cys-59, and polynucleotides encoding such polypeptides. In particular, the present invention provides polypeptides comprising the amino acid sequence of residues n-207 of SEQ ID NO:2, where n is an integer in the range of 1-59 and where Cys-59 is the position of the first residue from the N-terminus of the complete KDI polypeptide (shown in SEQ ID NO:2) believed to be required for activity of the KDI protein.

More in particular, the invention provides polypeptides having the amino acid sequence of residues 1-207, 2-207, 3-207, 4-207, 5-207, 6-207, 7-207, 8-207, 9-207, 10-207, 11-207, 12-207, 13-207, 14-207, 15-207, 16-207, 17-207, 18-207, 19-207, 20-207, 21-207, 22-207, 23-207, 24-207, 25-207, 26-207, 27-207, 28-207, 29-207, 30-207, 31-207, 32-207, 33-207, 34-207, 35-207, 36-207, 37-207, 38-207, 39-207, 40-207, 41-207, 42-207, 43-207, 44-207, 45-207, 46-207, 47-207, 48-207, 49-207, 50-207, 51-207, 52-207, 53-207, 54-207, 55-207, 56-207, 57-207, 58-207 and 59-207, all of SEQ ID NO:2. Polynucleotides encoding these polypeptides also are provided.

Similarly, many examples of biologically functional C-terminal deletion muteins are known. For instance, Interferon gamma shows up to ten times higher activities by deleting 8-10 amino acid residues from the carboxy terminus of the protein (Döbeli et al., *J. Biotechnology* 7:199-216 (1988). In the present case, since the protein of the invention is a member of the interferon polypeptide family, deletions of C-terminal amino acids up to the tryptophan residue at position 183 (W-183) in SEQ ID NO:2 may retain some biological activity such as antiviral activity or inhibition of bone marrow proliferation. Polypeptides having further

28-205, 29-205, 30-205, 31-205, 32-205, 33-205, 34-205, 35-205, 36-205, 37-205, 38-205, 39-205, 40-205, 41-205, 42-205, 43-205, 44-205, 45-205, 46-205, 47-205, 48-205, 49-205, 50-205, 51-205, 52-205, 53-205, 54-205, 55-205, 56-205, 57-205, 58-205, 59-205, 20-206, 21-206, 22-206, 23-206, 24-206, 25-206, 26-206, 27-206, 28-206, 29-206, 30-206, 31-206, 32-206, 33-206, 34-206, 35-206, 36-206, 37-206, 38-206, 39-206, 40-206, 41-206, 42-206, 43-206, 44-206, 45-206, 46-206, 47-206, 48-206, 49-206, 50-206, 51-206, 52-206, 53-206, 54-206, 55-206, 56-206, 57-206, 58-206 and 59-206 of SEQ ID NO:2. Each of the foregoing polypeptides may additionally include an N-terminal methionine residue. Polynucleotides encoding each of these polypeptides, with or without an N-terminal methionine residues are also are provided.

Also included are polypeptides consisting of a portion of the complete KDI amino acid sequence encoded by the human cDNA in clone HKAPI15, where this portion excludes from 1 to about 58 amino acids from the amino terminus of the complete amino acid sequence encoded by the human cDNA in clone HKAPI15, or from 1 to about 23 amino acids from the carboxy terminus, or any combination of the above amino terminal and carboxy terminal deletions, of the complete amino acid sequence encoded by the human cDNA in clone HKAPI15. Polynucleotides encoding all of the above deletion mutant polypeptide forms also are provided.

Other Mutants

In addition to terminal deletion forms of the protein discussed above, it also will be recognized by one of ordinary skill in the art that some amino acid sequences of the KDI polypeptide can be varied without significant effect on the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine activity.

Thus, the invention further includes variations of the KDI polypeptide which show substantial KDI polypeptide activity or which include regions of KDI protein such as the protein portions discussed below. Such mutants include deletions, insertions, inversions, repeats, and type substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," *Science 247*:1306-1310 (1990), wherein the authors indicate that there are two main approaches for studying the tolerance of an amino acid sequence to change. The first method relies on the process of evolution, in which mutations are either accepted or rejected by natural selection. The second approach uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene and selections or screens to identify sequences that maintain functionality.

C-terminal deletions including Ile-183 of SEQ ID NO:2 would not be expected to retain such biological activities because it is known that this residue in an interferon-related polypeptide is conserved among many members and is thought to be important for receptor binding and signal transduction. Furthermore, the cysteine residue at position 181 is highly conserved and known to be required for antiviral activity of members of the interferon family.

However, even if deletion of one or more amino acids from the C-terminus of a protein results in modification of loss of one or more biological functions of the protein, other biological activities may still be retained. Thus, the ability of the shortened protein to induce and/or bind to antibodies which recognize the complete protein generally will be retained when less than the majority of the residues of the complete protein are removed from the C-terminus. Whether a particular polypeptide lacking C-terminal residues of a complete protein retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art.

Accordingly, the present invention further provides polypeptides having one or more residues from the carboxy terminus of the amino acid sequence of the KDI shown in SEQ ID NO:2, up to Trp-183 of SEQ ID NO:2, and polynucleotides encoding such polypeptides. In particular, the present invention provides polypeptides having the amino acid sequence of residues 1-m of the amino acid sequence in SEQ ID NO:2, where m is any integer in the range of 183-207 and residue Trp-183 is the position of the first residue from the C- terminus of the complete KDI polypeptide (shown in SEQ ID NO:2) believed to be required for activity of the KDI protein.

More in particular, the invention provides polypeptides having the amino acid sequence of residues 1-183, 1-184, 1-185, 1-186, 1-187, 1-188, 1-189, 1-190, 1-191, 1-192, 1-193, 1-194, 1-195, 1-196, 1-197, 1-198, 1-199, 1-200, 1-201, 1-202, 1-203, 1-204, 1-205, 1-206 and 1-207 of SEQ ID NO:2. Polynucleotides encoding these polypeptides also are provided.

The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini, which may be described generally as having residues n-m of SEQ ID NO:2, where n and m are integers as described above. Furthermore, the invention provides these mutant polypeptides optionally having an N-terminal methionine. The polypeptides may therefore also be described by the formula x-n-m where X is either NH_2 or Met and n and m are integers as described above. Polynucleotides encoding these polypeptides are, of course, also provided.

More in particular, the invention preferrably provides polypeptides having the amino acid sequence of residues: 20-183, 21-183, 22-183, 23-183, 24-183, 25-183, 26-183, 27-183, 28-183, 29-183, 30-183, 31-183, 32-183, 33-183, 34-183, 35-183, 36-183, 37-183, 38-183, 39-183, 40-183, 41-183, 42-183, 43-183, 44-183, 45-183, 46-183, 47-183, 48-183, 49-183, 50-183, 51-183, 52-183, 53-183, 54-183, 55-183, 56-183, 57-183, 58-183, 59-183, 20-184, 21-184, 22-184, 23-184, 24-184, 25-184, 26-184, 27-184, 28-184, 29-184, 30-184, 31-184,

32-184, 33-184, 34-184, 35-184, 36-184, 37-184, 38-184, 39-184, 40-184, 41-184, 42-184, 43-184, 44-184, 45-184, 46-184, 47-184, 48-184, 49-184, 50-184, 51-184, 52-184, 53-184, 54-184, 55-184, 56-184, 57-184, 58-184, 59-184, 20-185, 21-185, 22-185, 23-185, 24-185, 25-185, 26-185, 27-185, 28-185, 29-185, 30-185, 31-185, 32-185, 33-185, 34-185, 35-185, 36-185, 37-185, 38-185, 39-185, 40-185, 41-185, 42-185, 43-185, 44-185, 45-185, 46-185, 47-185, 48-185, 49-185, 50-185, 51-185, 52-185, 53-185, 54-185, 55-185, 56-185, 57-185, 58-185, 59-185, 20-186, 21-186, 22-186, 23-186, 24-186, 25-186, 26-186, 27-186, 28-186, 29-186, 30-186, 31-186, 32-186, 33-186, 34-186, 35-186, 36-186, 37-186, 38-186, 39-186, 40-186, 41-186, 42-186, 43-186, 44-186, 45-186, 46-186, 47-186, 48-186, 49-186, 50-186, 51-186, 52-186, 53-186, 54-186, 55-186, 56-186, 57-186, 58-186, 59-186, 20-187, 21-187, 22-187, 23-187, 24-187, 25-187, 26-187, 27-187, 28-187, 29-187, 30-187, 31-187, 32-187, 33-187, 34-187, 35-187, 36-187, 37-187, 38-187, 39-187, 40-187, 41-187, 42-187, 43-187, 44-187, 45-187, 46-187, 47-187, 48-187, 49-187, 50-187, 51-187, 52-187, 53-187, 54-187, 55-187, 56-187, 57-187, 58-187, 59-187, 20-188, 21-188, 22-188, 23-188, 24-188, 25-188, 26-188, 27-188, 28-188, 29-188, 30-188, 31-188, 32-188, 33-188, 34-188, 35-188, 36-188, 37-188, 38-188, 39-188, 40-188, 41-188, 42-188, 43-188, 44-188, 45-188, 46-188, 47-188, 48-188, 49-188, 50-188, 51-188, 52-188, 53-188, 54-188, 55-188, 56-188, 57-188, 58-188, 59-188, 20-189, 21-189, 22-189, 23-189, 24-189, 25-189, 26-189, 27-189, 28-189, 29-189, 30-189, 31-189, 32-189, 33-189, 34-189, 35-189, 36-189, 37-189, 38-189, 39-189, 40-189, 41-189, 42-189, 43-189, 44-189, 45-189, 46-189, 47-189, 48-189, 49-189, 50-189, 51-189, 52-189, 53-189, 54-189, 55-189, 56-189, 57-189, 58-189, 59-189, 20-190, 21-190, 22-190, 23-190, 24-190, 25-190, 26-190, 27-190, 28-190, 29-190, 30-190, 31-190, 32-190, 33-190, 34-190, 35-190, 36-190, 37-190, 38-190, 39-190, 40-190, 41-190, 42-190, 43-190, 44-190, 45-190, 46-190, 47-190, 48-190, 49-190, 50-190, 51-190, 52-190, 53-190, 54-190, 55-190, 56-190, 57-190, 58-190, 59-190, 20-191, 21-191, 22-191, 23-191, 24-191, 25-191, 26-191, 27-191, 28-191, 29-191, 30-191, 31-191, 32-191, 33-191, 34-191, 35-191, 36-191, 37-191, 38-191, 39-191, 40-191, 41-191, 42-191, 43-191, 44-191, 45-191, 46-191, 47-191, 48-191, 49-191, 50-191, 51-191, 52-191, 53-191, 54-191, 55-191, 56-191, 57-191, 58-191, 59-191, 20-192, 21-192, 22-192, 23-192, 24-192, 25-192, 26-192, 27-192, 28-192, 29-192, 30-192, 31-192, 32-192, 33-192, 34-192, 35-192, 36-192, 37-192, 38-192, 39-192, 40-192, 41-192, 42-192, 43-192, 44-192, 45-192, 46-192, 47-192, 48-192, 49-192, 50-192, 51-192, 52-192, 53-192, 54-192, 55-192, 56-192, 57-192, 58-192, 59-192, 20-193, 21-193, 22-193, 23-193, 24-193, 25-193, 26-193, 27-193, 28-193, 29-193, 30-193, 31-193, 32-193, 33-193, 34-193, 35-193, 36-193, 37-193, 38-193, 39-193, 40-193, 41-193, 42-193, 43-193, 44-193, 45-193, 46-193, 47-193, 48-193, 49-193, 50-193, 51-193, 52-193, 53-193, 54-193, 55-193, 56-193, 57-193, 58-193, 59-193, 20-194, 21-194, 22-194, 23-194, 24-194, 25-194, 26-194, 27-194, 28-194, 29-194, 30-194, 31-194, 32-194, 33-194, 34-194, 35-194, 36-194, 37-194, 38-194, 39-194, 40-194, 41-194, 42-194, 43-194, 44-194, 45-194, 46-194, 47-194, 48-194, 49-194,

50-194, 51-194, 52-194, 53-194, 54-194, 55-194, 56-194, 57-194, 58-194, 59-194, 20-195, 21-195, 22-195, 23-195, 24-195, 25-195, 26-195, 27-195, 28-195, 29-195, 30-195, 31-195, 32-195, 33-195, 34-195, 35-195, 36-195, 37-195, 38-195, 39-195, 40-195, 41-195, 42-195, 43-195, 44-195, 45-195, 46-195, 47-195, 48-195, 49-195, 50-195, 51-195, 52-195, 53-195, 54-195, 55-195, 56-195, 57-195, 58-195, 59-195, 20-196, 21-196, 22-196, 23-196, 24-196, 25-196, 26-196, 27-196, 28-196, 29-196, 30-196, 31-196, 32-196, 33-196, 34-196, 35-196, 36-196, 37-196, 38-196, 39-196, 40-196, 41-196, 42-196, 43-196, 44-196, 45-196, 46-196, 47-196, 48-196, 49-196, 50-196, 51-196, 52-196, 53-196, 54-196, 55-196, 56-196, 57-196, 58-196, 59-196, 20-197, 21-197, 22-197, 23-197, 24-197, 25-197, 26-197, 27-197, 28-197, 29-197, 30-197, 31-197, 32-197, 33-197, 34-197, 35-197, 36-197, 37-197, 38-197, 39-197, 40-197, 41-197, 42-197, 43-197, 44-197, 45-197, 46-197, 47-197, 48-197, 49-197, 50-197, 51-197, 52-197, 53-197, 54-197, 55-197, 56-197, 57-197, 58-197, 59-197, 20-198, 21-198, 22-198, 23-198, 24-198, 25-198, 26-198, 27-198, 28-198, 29-198, 30-198, 31-198, 32-198, 33-198, 34-198, 35-198, 36-198, 37-198, 38-198, 39-198, 40-198, 41-198, 42-198, 43-198, 44-198, 45-198, 46-198, 47-198, 48-198, 49-198, 50-198, 51-198, 52-198, 53-198, 54-198, 55-198, 56-198, 57-198, 58-198, 59-198, 20-199, 21-199, 22-199, 23-199, 24-199, 25-199, 26-199, 27-199, 28-199, 29-199, 30-199, 31-199, 32-199, 33-199, 34-199, 35-199, 36-199, 37-199, 38-199, 39-199, 40-199, 41-199, 42-199, 43-199, 44-199, 45-199, 46-199, 47-199, 48-199, 49-199, 50-199, 51-199, 52-199, 53-199, 54-199, 55-199, 56-199, 57-199, 58-199, 59-199, 20-200, 21-200, 22-200, 23-200, 24-200, 25-200, 26-200, 27-200, 28-200, 29-200, 30-200, 31-200, 32-200, 33-200, 34-200, 35-200, 36-200, 37-200, 38-200, 39-200, 40-200, 41-200, 42-200, 43-200, 44-200, 45-200, 46-200, 47-200, 48-200, 49-200, 50-200, 51-200, 52-200, 53-200, 54-200, 55-200, 56-200, 57-200, 58-200, 59-200, 20-201, 21-201, 22-201, 23-201, 24-201, 25-201, 26-201, 27-201, 28-201, 29-201, 30-201, 31-201, 32-201, 33-201, 34-201, 35-201, 36-201, 37-201, 38-201, 39-201, 40-201, 41-201, 42-201, 43-201, 44-201, 45-201, 46-201, 47-201, 48-201, 49-201, 50-201, 51-201, 52-201, 53-201, 54-201, 55-201, 56-201, 57-201, 58-201, 59-201, 20-202, 21-202, 22-202, 23-202, 24-202, 25-202, 26-202, 27-202, 28-202, 29-202, 30-202, 31-202, 32-202, 33-202, 34-202, 35-202, 36-202, 37-202, 38-202, 39-202, 40-202, 41-202, 42-202, 43-202, 44-202, 45-202, 46-202, 47-202, 48-202, 49-202, 50-202, 51-202, 52-202, 53-202, 54-202, 55-202, 56-202, 57-202, 58-202, 59-202, 20-203, 21-203, 22-203, 23-203, 24-203, 25-203, 26-203, 27-203, 28-203, 29-203, 30-203, 31-203, 32-203, 33-203, 34-203, 35-203, 36-203, 37-203, 38-203, 39-203, 40-203, 41-203, 42-203, 43-203, 44-203, 45-203, 46-203, 47-203, 48-203, 49-203, 50-203, 51-203, 52-203, 53-203, 54-203, 55-203, 56-203, 57-203, 58-203, 59-203, 20-204, 21-204, 22-204, 23-204, 24-204, 25-204, 26-204, 27-204, 28-204, 29-204, 30-204, 31-204, 32-204, 33-204, 34-204, 35-204, 36-204, 37-204, 38-204, 39-204, 40-204, 41-204, 42-204, 43-204, 44-204, 45-204, 46-204, 47-204, 48-204, 49-204, 50-204, 51-204, 52-204, 53-204, 54-204, 55-204, 56-204, 57-204, 58-204, 59-204, 20-205, 21-205, 22-205, 23-205, 24-205, 25-205, 26-205, 27-205,

As the authors state, these studies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at a certain position of the protein. For example, most buried amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Other such phenotypically silent substitutions are described in Bowie, J. U. *et al., supra*, and the references cited therein. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

Thus, the fragment, derivative or analog of the polypeptide of SEQ ID NO:2, or that encoded by the deposited cDNA, may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the KDI polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the above form of the polypeptide, such as an IgG Fc fusion region peptide or leader or secretory sequence or a sequence which is employed for purification of the above form of the polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

Thus, the KDI of the present invention may include one or more amino acid substitutions, deletions or additions, either from natural mutations or human manipulation. As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein.

Amino acids in the KDI protein of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, *Science 244*:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as receptor binding or *in vitro* proliferative activity.

Of special interest are substitutions of charged amino acids with other charged or neutral amino acids which may produce proteins with highly desirable improved characteristics, such as less aggregation. Aggregation may not only reduce activity but also be problematic when preparing pharmaceutical formulations, because aggregates can be immunogenic (Pinckard *et al., Clin. Exp. Immunol. 2*:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland *et al., Crit. Rev. Therapeutic Drug Carrier Systems 10*:307-377 (1993).

Replacement of amino acids can also change the selectivity of the binding of a ligand to cell surface receptors. For example, Ostade *et al., Nature 361*:266-268 (1993) describes certain mutations resulting in selective binding of TNF-α to only one of the two known types of TNF receptors. Sites that are critical for ligand-receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al., J. Mol. Biol. 224*:899-904 (1992) and de Vos *et al. Science 255*:306-312 (1992)).

Particularly preferred substitutions for each of the KDI polypeptides described herein is the replacement of the arginine residues at position 192 with lysine (sometimes hereinafter referred to as "R192K"), and replacement of the cysteine residue at position 193 with a serine residue (sometimes hereinafter referred to as "C193S"). These substitutions can be found in a KDI polypeptide indivudually or they can occur in the same KDI polypeptide. Polynucleotides encoding all of the foregoing KDI polypeptides containing substitutions.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of the KDI polypeptide can be substantially purified by the one-step method described in Smith and Johnson, *Gene 67*:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using anti-KDI antibodies of the invention in methods which are well known in the art of protein purification.

The invention further provides an isolated KDI polypeptide comprising an amino acid sequence selected from the group consisting of: (a) the amino acid sequence of the full-length KDI polypeptide having the complete amino acid sequence shown in SEQ ID NO:2; (b) the amino acid sequence of the full-length KDI polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 excepting the N-terminal methionine (i.e., residues 2 to 207 of SEQ ID NO:2); the amino acid sequence of the mature KDI polypeptide shown as residues 28-207 in SEQ ID NO:2; (d) the amino acid sequence shown in SEQ ID NO:2 as residues 165-183; (e) the full length KDI polypeptide encoded by the human cDNA contained in clone HKAPI15; (f) the full-length KDI polypeptide encoded by the human cDNA contained in clone HKAPI15 excepting the N-terminal methionine; and (g) the mature KDI polypeptide encoded by the human cDNA contained in clone HKAPI15.

Further polypeptides of the present invention include polypeptides which have at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98% or 99% similarity to those described above. The polypeptides of the invention also comprise those which are at least 80% identical, more preferably at least 90% or 95% identical, still more preferably at least 96%, 97%, 98% or 99% identical to the polypeptide encoded by the deposited DNA or to the polypeptide of SEQ ID NO:2, and also include portions of such polypeptides with at least 10, 20 or 30 amino acids and more preferably at least 50 amino acids.

A further embodiment of the invention relates to a peptide or polypeptide which comprises the amino acid sequence of a KDI polypeptide having an amino acid sequence which

contains at least one conservative amino acid substitution, but not more than 50 conservative amino acid substitutions, even more preferably, not more than 40 conservative amino acid substitutions, still more preferably not more than 30 conservative amino acid substitutions, and still even more preferably not more than 20 conservative amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a peptide or polypeptide to have an amino acid sequence which comprises the amino acid sequence of a KDI polypeptide, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 conservative amino acid substitutions.

By "% similarity" for two polypeptides is intended a similarity score produced by comparing the amino acid sequences of the two polypeptides using the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711) and the default settings for determining similarity. Bestfit uses the local homology algorithm of Smith and Waterman (Advances in Applied Mathematics 2:482-489, 1981) to find the best segment of similarity between two sequences.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a reference amino acid sequence of an IL-20 polypeptide is intended that the amino acid sequence of the polypeptide is identical to the query sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of the IL-20 polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the query sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence shown in SEQ ID NO:2 or to the amino acid sequence encoded by deposited cDNA clone can be determined conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference

sequence are allowed. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence

are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The polypeptide of the present invention could be used as a molecular weight marker on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art.

As described in detail below, the polypeptides of the present invention can also be used to raise polyclonal and monoclonal antibodies, which are useful in assays for detecting KDI protein expression as described below or as agonists and antagonists capable of enhancing or inhibiting KDI protein function. Further, such polypeptides can be used in the yeast two-hybrid system to "capture" KDI protein binding proteins which are also candidate agonists and antagonists according to the present invention. The yeast two hybrid system is described in Fields and Song, Nature 340:245-246 (1989).

Epitope-Bearing Portions

In another aspect, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes. See, for instance, Geysen *et al., Proc. Natl. Acad. Sci. USA 81*:3998- 4002 (1983).

As to the selection of peptides or polypeptides bearing an antigenic epitope (i.e., that contain a region of a protein molecule to which an antibody can bind), it is well known in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein. See, for instance, Sutcliffe, J. G., Shinnick, T. M., Green, N. and Learner, R. A. (1983) "Antibodies that react with predetermined sites on proteins," *Science, 219*:660-666. Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins (i.e., immunogenic epitopes) nor to the amino or carboxyl terminals. Antigenic epitope-bearing peptides and polypeptides of the invention are therefore useful to raise antibodies, including monoclonal antibodies, that bind specifically to a polypeptide of the invention. See, for instance, Wilson *et al., Cell 37*:767-778 (1984) at 777.

Antigenic epitope-bearing peptides and polypeptides of the invention preferably contain a sequence of at least seven, more preferably at least nine and most preferably between about 15 to about 30 amino acids contained within the amino acid sequence of a polypeptide of the invention. Non-limiting examples of antigenic polypeptides or peptides that can be used to generate KDI-specific antibodies include: a polypeptide comprising amino acid residues from

about Ser 49 to about Ser 54 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about Cys 59 to about Ala 65 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about Pro 78 to about Tyr 88 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about His 101 to about Gln 113 in SEQ ID NO:2; a polypeptide comprising amino acid residues Gln 120 to about Glu 123 in SEQ ID NO:2; a polypeptide comprising amino acid residues Cys 128 to about Pro 155 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Leu 160 to about Arg 168 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about Asn 171 to about Asp 180 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about Val 186 to about Cys 193 in SEQ ID NO:2; and a polypeptide comprising amino acid residues Phe 204 to about Lys 207 in SEQ ID NO:2. These polypeptide fragments have been determined to bear antigenic epitopes of the KDI protein by the analysis of the Jameson-Wolf antigenic index, as shown in Figure 3, above.

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means. See, e.g., Houghten, R. A. (1985) "General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids." *Proc. Natl. Acad. Sci. USA 82*:5131-5135; this "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten *et al.* (1986).

Epitope-bearing peptides and polypeptides of the invention are used to induce antibodies according to methods well known in the art. See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. *et al., Proc. Natl. Acad. Sci. USA 82*:910-914; and Bittle, F. J. *et al., J. Gen. Virol. 66*:2347-2354 (1985). Immunogenic epitope-bearing peptides of the invention, i.e., those parts of a protein that elicit an antibody response when the whole protein is the immunogen, are identified according to methods known in the art. See, for instance, Geysen et al., *supra*. Further still, U.S. Patent No. 5,194,392 to Geysen (1990) describes a general method of detecting or determining the sequence of monomers (amino acids or other compounds) which is a topological equivalent of the epitope (i.e., a "mimotope") which is complementary to a particular paratope (antigen binding site) of an antibody of interest. More generally, U.S. Patent No. 4,433,092 to Geysen (1989) describes a method of detecting or determining a sequence of monomers which is a topographical equivalent of a ligand which is complementary to the ligand binding site of a particular receptor of interest. Similarly, U.S. Patent No. 5,480,971 to Houghten, R. A. et al. (1996) on Peralkylated Oligopeptide Mixtures discloses linear C1-C7-alkyl peralkylated oligopeptides and sets and libraries of such peptides, as well as methods for using such oligopeptide sets and libraries for determining the sequence of a peralkylated oligopeptide that preferentially binds to an acceptor molecule of interest. Thus, non-peptide analogs of the epitope-bearing peptides of the invention also can be made routinely by these methods.

Fusion Proteins

As one of skill in the art will appreciate, KDI polypeptides of the present invention and the epitope-bearing fragments thereof described above can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. This has been shown, e.g., for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988)). Fusion proteins that have a disulfide-linked dimeric structure due to the IgG part can also be more efficient in binding and neutralizing other molecules than the monomeric KDI protein or protein fragment alone (Fountoulakis et *al., J. Biochem. 270*:3958-3964 (1995)).

Antibodies

KDI-protein specific antibodies for use in the present invention can be raised against the intact KDI protein or an antigenic polypeptide fragment thereof, which may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier.

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to KDI protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding of an intact antibody (Wahl *et al., J. Nucl. Med. 24*:316-325 (1983)). Thus, these fragments are preferred.

The antibodies of the present invention may be prepared by any of a variety of methods. For example, cells expressing the KDI protein or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of KDI protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or KDI protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology (Köhler *et al., Nature 256*:495 (1975); Köhler *et al., Eur. J. Immunol. 6*:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling *et al., in*: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., (1981) pp. 563-681). In general, such procedures involve immunizing an animal (preferably a mouse) with a KDI protein antigen or, more preferably, with a KDI protein-expressing cell. Suitable cells can be recognized by their capacity to bind anti-KDI protein antibody. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56° C), and

supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 μg/ml of streptomycin. The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the American Type Culture Collection, Rockville, Maryland. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the KDI protein antigen.

Alternatively, additional antibodies capable of binding to the KDI protein antigen may be produced in a two-step procedure through the use of anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and that, therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, KDI-protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the KDI protein-specific antibody can be blocked by the KDI protein antigen. Such antibodies comprise anti-idiotypic antibodies to the KDI protein-specific antibody and can be used to immunize an animal to induce formation of further KDI protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, KDI protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For *in vivo* use of anti-KDI in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. See, for review, Morrison, *Science* *229*:1202 (1985); Oi et al., *BioTechniques 4*:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne *et al., Nature 312*:643 (1984); Neuberger *et al., Nature 314*:268 (1985).

Immune System-Related Disorders

Treatment

It will also be appreciated by one of ordinary skill that, since the KDI protein of the invention is a member of the interferon family, when KDI is added to cells, tissues or the body

of an individual, the protein will exert its physiological activities on its target cells of that individual. Therefore, it will be appreciated that conditions caused by a decrease in the standard or normal level of interferon activity in an individual, particularly disorders of the immune system, can be treated by administration of the KDI polypeptide. Thus, the invention also provides a method of treatment of an individual in need of an increased level of interferon activity comprising administering to such an individual a pharmaceutical composition comprising an amount of an isolated KDI polypeptide of the invention, effective to increase the interferon activity level in such an individual.

The human class I IFN receptor complex which mediates the biological activity of IFN-alpha and IFN-beta also binds IFN-omega and is expected to bind KDI. Accordingly, KDI can be used clinically for anti-viral therapy, for example, in the treatment of AIDS, viral hepatitis including chronic hepatitis B, hepatitis C, papilloma viruses, viral encephalitis, and in the prophylaxis of rhinitis and respiratory infections.

KDI is also useful in the treatment of numerous cancers (*e.g.*, hairy cell leukemia, acute myeloid leukemia, osteosarcoma, basal cell carcinoma, glioma, renal cell carcinoma, multiple myeloma, melanoma, and Hodgkin's disease).

KDI is believed to stimulate natural killer cell activity. Accordingly, KDI may be used to treat parasitic and bacterial infection for example, for treating *Cryptosporidium parvum* infection and multidrug-resistant pulmonary tuberculosis.

KDI is also believed to be useful as an immunotherapeutic agent, more specifically as an immunosuppressive agent. For example, KDI is believed to inhibit proliferation of lymphocytes stimulated with mitogens or allogeneic cells, myeloid progenitor cells and other bone marrow cells. Accordingly, KDI is useful as a protective agent when administered prior to chemotherapy and in addition can be used to treat hyperproliferation of lymphocytes, myeloid progenitors and bone marrow stem cells, e.g., in the treatment of chronic myelogenous leukemia. KDI polypeptides can also be used in the prevention of graft vs. host rejection, or to curtail the progresion of autoimmune diseases, such as arthritis, multiple sclerosis, (2) or diabetes (3). KDI is also useful in the treatment of allergies in mammals, e.g., by inhibiting the humoral response.

KDI may be used as an adjuvant or coadjuvant to enhance or simulate the immune response in cases of prophylactic or therapeutic vaccination.

Further, there is provided a method of treating infection in a patient comprising administering an effective amount of a polypeptide of the invention to a patient in need of anti-infective therapy. In a preferred embodiment the infection is of viral, bacterial, or parasitic etiology. In a particularly preferred embodiment, the infection is a viral infection.

Further, there is provided a method of treating cancer in a patient comprising administerin an effective amount of a polypeptide of the invention to a patient in need of anti-cancer therapy.

Further, there is provided a method of immunotherapy in a patient comprising administering an effective amount of a polypeptide of the invention to a patient in need of immunotherapy.

Formulations

The KDI polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with KDI polypeptide alone), the site of delivery of the KDI polypeptide composition, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" of KDI polypeptide for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of KDI polypeptide administered parenterally per dose will be in the range of about 1 μg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the KDI polypeptide is typically administered at a dose rate of about 1 μg/kg/hour to about 50 μg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the KDI of the invention may be administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments, drops or transdermal patch), bucally, or as an oral or nasal spray. By "pharmaceutically acceptable carrier" is meant a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The KDI polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., *Biopolymers 22*:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., *J. Biomed. Mater. Res. 15*:167-277 (1981), and R. Langer, *Chem. Tech. 12*:98-105 (1982)), ethylene vinyl acetate (R. Langer et al., Id.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release KDI polypeptide compositions also include liposomally entrapped KDI polypeptide. Liposomes containing KDI polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl.

Acad. Sci. (USA) 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. (USA) 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal KDI polypeptide therapy.

For parenteral administration, in one embodiment, the KDI polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the KDI polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The KDI polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of KDI polypeptide salts.

KDI polypeptide to be used for therapeutic administration must be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic KDI polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

KDI polypeptide ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous KDI polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized KDI polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Agonists and Antagonists - Assays and Molecules

The invention also provides a method of screening compounds to identify those which enhance or block the action of KDI on cells, such as its interaction with KDI-binding molecules such as receptor molecules. An agonist is a compound which increases the natural biological functions of KDI or which functions in a manner similar to KDI, while antagonists decrease or eliminate such functions.

In another aspect of this embodiment the invention provides a method for identifying a receptor protein or other ligand-binding protein which binds specifically to a KDI polypeptide. For example, a cellular compartment, such as a membrane or a preparation thereof, may be prepared from a cell that expresses a molecule that binds KDI. The preparation is incubated with labeled KDI. KDI and complexes of KDI bound to the receptor or other binding protein are isolated and characterized according to routine methods known in the art. Alternatively, the KDI polypeptide may be bound to a solid support so that binding molecules solubilized from cells are bound to the column and then eluted and characterized according to routine methods.

In the assay of the invention for agonists or antagonists, a cellular compartment, such as a membrane or a preparation thereof, may be prepared from a cell that expresses a molecule that binds KDI, such as a molecule of a signaling or regulatory pathway modulated by KDI. The preparation is incubated with labeled KDI in the absence or the presence of a candidate molecule which may be a KDI agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labeled ligand. Molecules which bind gratuitously, i.e., without inducing the effects of KDI on binding the KDI binding molecule, are most likely to be good antagonists. Molecules that bind well and elicit effects that are the same as or closely related to KDI are agonists.

KDI-like effects of potential agonists and antagonists may by measured, for instance, by determining activity of a second messenger system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of KDI or molecules that elicit the same effects as KDI. Second messenger systems that may be useful in this regard include but are not limited to AMP guanylate cyclase, ion channel or phosphoinositide hydrolysis second messenger systems.

Another example of an assay for KDI antagonists is a competitive assay that combines KDI and a potential antagonist with membrane-bound KDI receptor molecules or recombinant KDI receptor molecules under appropriate conditions for a competitive inhibition assay. KDI can be labeled, such as by radioactivity, such that the number of KDI molecules bound to a receptor molecule can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a receptor molecule, without inducing KDI-induced activities, thereby preventing the action of KDI by excluding KDI from binding.

Other potential antagonists include antisense molecules. Antisense technology can be used to control gene expression through antisense DNA or RNA or through triple-helix formation. Antisense techniques are discussed, for example, in Okano, *J. Neurochem. 56*: 560 (1991); "Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression." CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance *Lee et al., Nucleic Acids Research 6*: 3073 (1979); Cooney *et al., Science 241*: 456 (1988); and *Dervan et al., Science 251*: 1360 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA. For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of KDI. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into KDI polypeptide. The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of KDI protein.

The agonists and antagonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as described above.

The antagonists may be employed for instance to inhibit interferon activity, for example, following chemotherapy to stimulate proliferation of bone marrow and haematopoietic

progenitor cells. Any of the above antagonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as hereinafter described.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1: Cloning and Expression of KDI in E. coli

The novel pHE4 series of bacterial expression vectors, in particular, the pHE4a vector is used for bacterial expression in this example. (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311). pHE4-5/KDI vector plasmid DNA contains the KDI coding polynucleotide shown in Figure 1 inserted between unique restriction enzyme sites *NdeI* and *Asp718*. The construct was deposited with the ATCC on February 25, 1998 and given Accession No. 209645, as a convenience to those of skill in the art.

The pHE4a bacterial expression vector includes a neomycin phosphotransferase gene for selection, an *E. coli* origin of replication, a T5 phage promoter sequence, two *lac* operator sequences, a Shine-Delgarno sequence, and the lactose operon repressor gene (*lac*Iq). These elements are arranged such that an inserted DNA fragment encoding a polypeptide expresses that polypeptide with the six His residues (i.e., a "6 X His tag") covalently linked to the amino terminus of that polypeptide.

The DNA sequence encoding the mature KDI protein is amplified using PCR oligonucleotide primers which anneal to the amino terminal sequences of the desired portion of the KDI protein and to sequences in the deposited construct 3' to the cDNA coding sequence. Additional nucleotides containing restriction sites to facilitate cloning in the pHE4a vector are added to the 5' and 3' primer sequences, respectively.

For cloning the KDI protein coding region, the 5' primer has the sequence 5' GGCCGCATATGCTGGACTGTAACTTACTG 3' (SEQ ID NO:16) containing the underlined NdeI restriction site. One of ordinary skill in the art would appreciate, of course, that the point in the protein coding sequence where the 5' primer begins may be varied to amplify a DNA segment encoding any desired portion of the complete KDI protein. The 3' primer has the sequence 5' GGCCGCGGTACCTTATTTCCTCCTGAATAGAGC 3' (SEQ ID NO:17) containing the underlined Asp718 restriction site.

The amplified KDI DNA fragment is digested with NdeI and Asp718 and it and the linearized plasmid are then ligated together. Insertion of the KDI DNA into the restricted pHE4a vector places the KDI protein coding region downstream from the IPTG-inducible promoter and in-frame with an initiating AUG and the six histidine codons.

The ligation mixture is transformed into competent *E. coli* cells using standard procedures such as those described by Sambrook and colleagues (*Molecular Cloning: a Laboratory Manual, 2nd Ed.*; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)). *E. coli* strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses the lac repressor and confers kanamycin resistance ("Kanr"), is used in carrying out the illustrative example described herein. This strain, which is only one of many that are suitable for expressing KDI protein, is available commercially (QIAGEN, Inc., *supra*). Transformants are identified by their ability to grow on LB plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis, PCR and DNA sequencing.

Clones containing the desired constructs are grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 μg/ml) and kanamycin (25 μg/ml). The O/N culture is used to inoculate a large culture, at a dilution of approximately 1:25 to 1:250. The cells are grown to an optical density at 600 nm ("OD600") of between 0.4 and 0.6. Isopropyl-β-D-thiogalactopyranoside ("IPTG") is then added to a final concentration of 1 mM to induce transcription from the lac repressor sensitive promoter, by inactivating the lacI repressor. Cells subsequently are incubated further for 3 to 4 hours. Cells then are harvested by centrifugation.

The cells are then stirred for 3-4 hours at 4°C in 6M guanidine-HCl, pH 8. The cell debris is removed by centrifugation, and the supernatant containing the KDI polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist, 1995, QIAGEN, Inc., *supra*). Briefly the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the KDI is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins can be eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50

mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4° C or frozen at -80° C.

The following alternative method may be used to purify KDI expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells are harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells ware then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 x *g* for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 x *g* centrifugation for 15 min., the pellet is discarded and the KDI polypeptide-containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 x *g*) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded KDI polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 μm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 mm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the KDI polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the KDI polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant KDI polypeptide exhibits greater than 95% purity after the above refolding and purification steps. No major contaminant bands are observed from Commassie blue stained 16% SDS-PAGE gel when 5 μg of purified protein is loaded. The purified protein is also tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

The following alternative method may be used to purify KDI expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells are harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells ware then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 x *g* for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 x *g* centrifugation for 15 min., the pellet is discarded and the KDI polypeptide-containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 x *g*) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring.

The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded KDI polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 μm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 mm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the KDI polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the KDI polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant KDI polypeptide exhibits greater than 95% purity after the above refolding and purification steps. No major contaminant bands are observed from Commassie blue stained 16% SDS-PAGE gel when 5 μg of purified protein is loaded. The purified protein is also tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Multiple KDI expression constructs have been generated by the inventors herein to facilitate the production of KDI polypeptides of several sizes and in several systems. The E. coli based constucts are as follows: (1) pQE9:KDI.S27-K207 (expresses amino acids 27-207 of SEQ ID NO:2); (2) pHE4:KDI.S27-K207 (expresses amino acids 27-207 of SEQ ID NO:2); (3) pHE4:KDI.A23-K207 (expresses amino acids 23-207 of SEQ ID NO:2); (4) pHE4.KDI.G24-K207 (expresses amino acids 24-207 of SEQ ID NO:2); and (5) pHE4.KDI.C30-K207 (expresses amino acids 30-207 of SEQ ID NO:2).

Example 2: Cloning and Expression of KDI protein in a Baculovirus Expression System

In this illustrative example, the plasmid shuttle vector pA2 GP is used to insert the cloned DNA encoding KDI, into a baculovirus to express the KDI protein, using a baculovirus leader and standard methods as described in Summers et al., *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas Agricultural Experimental Station Bulletin No. 1555 (1987). This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by the secretory signal peptide (leader) of the baculovirus gp67 protein and convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate viable virus that expresses the cloned polynucleotide.

Many other baculovirus vectors could be used in place of the vector above, such as pAc373, pVL941 and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology 170*:31-39 (19989).

The cDNA sequence encoding the mature KDI protein in the deposited clone, lacking the AUG initiation codon and the naturally associated leader sequence shown in SEQ ID NO:2, is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene. The 5' primer has the sequence 5' GGCCGGGATCCGCCATCATGAGCACCAAACCTGATATG 3' (SEQ ID NO:18) containing the underlined BamHI restriction enzyme site. The 3' primer has the sequence 5' GGCCGCGGTACCTTATTTCCTCCTGAATAGAGC 3' (SEQ ID NO:19) containing the underlined Asp718 restriction site.

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with BamHI and Asp718 and again is purified on a 1% agarose gel.

The plasmid is digested with the restriction enzymes BamHI and Asp718 and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures

known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

Fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Statagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria are identified that contain the plasmid with the human KDI gene by digesting DNA from individual colonies using BamHI and Asp718 and then analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing. This plasmid is designated herein pA2GPKDI.

Five μg of the plasmid pA2GPKDI is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., *Proc. Natl. Acad. Sci. USA 84*: 7413-7417 (1987). One μg of BaculoGold™ virus DNA and 5 μg of the plasmid pA2GPKDI are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10). After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 μl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C. The recombinant virus is called V-KDI.

To verify the expression of the KDI gene Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-KDI at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus

methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μCi of ^{35}S-methionine and 5 μCi ^{35}S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the of the KDI protein.

Other baculovirus expression constructs were constructed as follows: (1) pA2:KDI (which expresses residues 1-207 of SEQ ID NO:2); (2) pA2.KDI.M7-K207 (which expresses residues 7-207 of SEQ ID NO:2); (3) pA2gp.KDI.L28-K207 (which expresses residues 28-207 of SEQ ID NO:2); (4) pA2gp.KDI.C30-K207 (which expresses residues 30-207 of SEQ ID NO:2); and (5) pA2.KDI.M1-R192 (which expresses residues 1 to 192 of SEQ ID NO:2).

Example 3: Cloning and Expression of KDI in Mammalian Cells

A typical mammalian expression vector contains the promoter element, which mediates the initiation of transcription of mRNA, the protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription can be achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter). Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) and pBC12MI (ATCC 67109). Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the gene can be expressed in stable cell lines that contain the gene integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful to develop cell lines that carry several hundred or even several thousand copies of the gene of interest. Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., *Biochem J. 227*:277-279 (1991); Bebbington et al., *Bio/Technology 10*:169-175 (1992)). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are

selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

The expression vectors pC1 and pC4 contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen *et al., Molecular and Cellular Biology,* 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (*Boshart et al., Cell 41*:521-530 (1985)). Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors contain in addition the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene.

Cloning and Expression in CHO Cells

The vector pC4-Sig is used for the expression of KDI polypeptide. Plasmid pC4-Sig is a derivative of the plasmid pSV2-dhfr (ATCC Accession No. 37146). It contains coding region for the secretory leader sequence from chemokine beta-8 (see US95/09508) upstream from the multiple cloning site and is designed to be inframe with inserted heterologous DNA. The plasmid contains the mouse DHFR gene under control of the SV40 early promoter. Chinese hamster ovary- or other cells lacking dihydrofolate activity that are transfected with these plasmids can be selected by growing the cells in a selective medium (alpha minus MEM, Life Technologies) supplemented with the chemotherapeutic agent methotrexate. The amplification of the DHFR genes in cells resistant to methotrexate (MTX) has been well documented (see, e.g., Alt, F. W., Kellems, R. M., Bertino, J. R., and Schimke, R. T., 1978, *J. Biol. Chem. 253*:1357-1370, Hamlin, J. L. and Ma, C. 1990, *Biochem. et Biophys. Acta, 1097*:107-143, Page, M. J. and Sydenham, M. A. 1991, *Biotechnology 9*:64-68). Cells grown in increasing concentrations of MTX develop resistance to the drug by overproducing the target enzyme, DHFR, as a result of amplification of the DHFR gene. If a second gene is linked to the DHFR gene, it is usually co-amplified and over-expressed. It is known in the art that this approach may be used to develop cell lines carrying more than 1,000 copies of the amplified gene(s). Subsequently, when the methotrexate is withdrawn, cell lines are obtained which contain the amplified gene integrated into one or more chromosome(s) of the host cell.

Plasmid pC4 contains for expressing the gene of interest the strong promoter of the long terminal repeat (LTR) of the Rouse Sarcoma Virus (Cullen, et al., *Molecular and Cellular Biology*, March 1985:438-447) plus a fragment isolated from the enhancer of the immediate early gene of human cytomegalovirus (CMV) (Boshart et al., *Cell 41*:521-530 (1985)). Downstream of the promoter are the following single restriction enzyme cleavage sites that allow the integration of the genes: BamHI, Xba I, and Asp718. Behind these cloning sites the plasmid contains the 3' intron and polyadenylation site of the rat preproinsulin gene. Other high efficiency promoters can also be used for the expression, e.g., the human ß-actin promoter, the SV40 early or late promoters or the long terminal repeats from other retroviruses,

e.g., HIV and HTLVI. Clontech's Tet-Off and Tet-On gene expression systems and similar systems can be used to express the KDI polypeptide in a regulated way in mammalian cells (Gossen, M., & Bujard, H. 1992, *Proc. Natl. Acad. Sci. USA 89*:5547-5551). For the polyadenylation of the mRNA other signals, e.g., from the human growth hormone or globin genes can be used as well. Stable cell lines carrying a gene of interest integrated into the chromosomes can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

The plasmid pC4 is digested with the restriction enzymes BamHI and Asp718 and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

The DNA sequence encoding the KDI polypeptide is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the desired portion of the gene. The 5' primer containing the underlined BamHI site, a Kozak sequence, and an AUG start codon, has the following sequence:
5' GGCCG<u>GGATCC</u>GCCATCATGAGCACCAAACCTGATATG 3' (SEQ ID NO:18). The 3' primer, containing the underlined Asp718 restriction site has the following sequence: 5' GGCCGC<u>GGTACC</u>TTATTTCCTCCTGAATAGAGC 3' (SEQ ID NO:19).

The amplified fragment is digested with the endonucleases BamHI and Asp718 and then purified again on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC4 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene are used for transfection. Five μg of the expression plasmid pC4 is cotransfected with 0.5 μg of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μM, 2 μM, 5 μM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μM. Expression of the desired

gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Other mammalian expression vectors constructed are as follows: (1) pC4:KDI (which expresses residues 1-207 of SEQ ID NO:2); (2) pC4sp:KDI.C30-K207 (which expresses a heterologous signal peptide (the chemokine beta-8 (MPIF-1) signal peptide) followed by amino acid residues 30-207 of KDI; and (3) pC4sp:KDI.L28-K207 (which expresses a heterologous signal peptide (the chemokine beta-8 (MPIF-1) signal peptide) followed by amino acid residues 28-207 of KDI.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of all publications (including patents, patent applications, journal articles, laboratory manuals, books, or other documents) cited herein are hereby incorporated by reference.